Observations on REM sleep without atonia assessed with a semi-automated scoring algorithm

Jesper Jeppesen[1], Marit Otto[1], Yoon Frederiksen[2], Allan K Hansen[3], Tatyana D Fedorova[3], Karoline Knudsen[3], Adjmal Nahimi[3], David J Brooks[3,4,5], Per Borghammer[3], and Michael Sommerauer[3,6]

[1]Department of Clinical Neurophysiology, Aarhus University Hospital, Aarhus, Denmark
[2]Department of Clinical Medicine & Department of Psychology, Aarhus University, Aarhus, Denmark
[3]Department of Nuclear Medicine and PET Centre, Aarhus University Hospital, Aarhus, Denmark
[4]Division of Neuroscience, Department of Medicine, Imperial College London, London, UK
[5]Division of Neuroscience, Newcastle University, Newcastle, UK
[6]Department of Neurology, University Hospital Cologne, Cologne, Germany

Corresponding author	Michael Sommerauer, Department of Nuclear Medicine and PET Centre, Aarhus University Hospital, Denmark, Noerrebrogade 44, Building 10G, 6th floor, 8000 Aarhus C; Michael.sommerauer@clin.au.dk
Financial support	This project received financial support from the Lundbeck Foundation, the Danish Association of Parkinson's disease (Parkinson-foreningen), the Swiss National Science Foundation (grant number, P2SKP3_161812), and the Hildegard Henssler-Stiftung.
Conflicts of interest	The authors declare no conflicts of interests with regards to this manuscript.

Highlights

- REM sleep behavior disorder includes failure of muscle atonia but its visual assessment is demanding.
- The proposed computer algorithm could quantify muscle activity comparable to human scorings and detected alterations in duration and amplitude of muscle activity.
- The computer-based scoring might be a useful tools for quantification of altered muscle activity and detection of REM sleep behavior disorder.

Abstract

Background: Rapid eye movement (REM) sleep behavior disorder (RBD) is defined by dream enactment and a failure of muscle atonia. Visual assessment of this muscle activity is time consuming and rater-dependent. Therefore, automated approaches are desired.

Methods: A computer algorithm for scoring of 'tonic', 'phasic' and 'any' muscle activity was evaluated compared with ratings from human observers. Subsequently, 52 subjects were analyzed with the algorithm. Duration and maximal amplitude of muscle activity, and self-awareness of RBD symptoms were also assessed.

Results: The computer algorithm showed high congruency with human ratings and all subjects with RBD were correctly identified by excess of 'tonic', 'phasic' and 'any' submental muscle activity, when artifacts were removed before analysis. Subjects with RBD exhibited prolonged 'phasic' muscle activity bouts with higher amplitude, and self-awareness of RBD symptoms correlated with amount of REM sleep without atonia.

Discussion: Our proposed algorithm was able to detect and rate REM sleep without atonia, and allowed identification of RBD. Increased duration and amplitude of muscle activity bouts were additional characteristics of RBD. Quantification of REM sleep without atonia might represent a marker of RBD severity.

Significance: Our computer algorithm can support diagnosis of RBD and quantification of altered muscle activity.

Keywords: REM sleep without atonia, REM sleep behavior disorder, Parkinson's disease

1. Introduction

Rapid eye movement (REM) sleep behavior disorder (RBD) is a disorder defined by inappropriate muscle activity and enactment of dream content during REM sleep. RBD is closely connected to diseases with pathological aggregation of α-synuclein, including Parkinson's disease (PD), dementia with Lewy bodies and multiple system atrophy (St Louis et al. , 2017). RBD often precedes overt motor symptoms of α-synucleinopathies by years (Postuma et al. , 2015b, St Louis et al. , 2017), providing a window to study the early pre-motor and prodromal disease stages. Additionally, occurrence of RBD in established PD is related to more rapid and severe disease progression with a higher burden of non-motor symptoms (Neikrug et al. , 2014, Chahine et al. , 2016). Visual quantifications of REM sleep without atonia (RSWA) subclassify altered muscle activity into 'tonic' activity, representing sustained muscle activity and 'phasic' activity representing short lasting activity (Lapierre et al. , 1992, Frauscher et al. , 2012, McCarter et al. , 2014). Additionally, the term 'any' activity was introduced to account for overall muscle activity (Frauscher et al. , 2012). However, visual scorings are demanding and time consuming (Frauscher et al. , 2013); hence, computer-assisted quantification is a desirable tool.

We developed a computer algorithm, which follows recently published rules for visual scoring of altered muscle activity proposed by McCarter and colleagues, which was reported to have high diagnostic accuracy even in the presence of sleep apnea (McCarter et al. , 2014). We integrated visual assessments of sleep stages, breathing events, arousals, and technical artifacts that are typically assessed during routine evaluation of a polysomnography (PSG), in this semi-automated algorithm. We aimed to assess comparability of scorings from the algorithm with two human raters and we subsequently tested the algorithm on 52 subjects. Additionally, we investigated if qualitative aspects of RSWA assessed with the algorithm - i.e. duration and amplitude of muscle activity - were different between subjects with and without RBD and if the amount of RSWA correlated to self-reported symptom awareness.

2. Methods

2.1. Subjects

We recruited 52 subjects, grouped into 23 subjects with RBD (RBD+), and 29 subjects without RBD (RBD-). The RBD- group included healthy subjects (healthy controls, HC; n = 15) and subjects with PD (PD^{RBD-}; n = 14); the RBD+ group included subjects with RBD but without signs of any other neurological disorder ('idiopathic' RBD, iRBD; n = 7), and subjects with PD (PD^{RBD+}; n = 16). The clinical diagnosis of PD followed published consensus criteria from the Movement Disorder Society (MDS) (Postuma et al. , 2015a). Basic demographic characteristics were assessed as well as RBD symptoms with the RBD screening questionnaire (RBDSQ) (Stiasny-Kolster et al. , 2007). No subjects were receiving antidepressants or benzodiazepines; melatonin was stopped two weeks prior to sleep examination in one subject. Motor symptoms of PD patients were scored with the MDS Unified PD Rating Scale part III (MDS-UPDRS III) after 12 hours of medication withdrawal. PD disease duration, dopaminergic medication and Hoehn and Yahr disease stage were also documented. Levodopa equivalent doses (LED) were calculated as proposed by Tomlinson et al (Tomlinson et al. , 2010). We did not record beginning of RBD as subjects reported to strong difficulties in reliable determination of the beginning of their RBD symptoms. All subjects were recruited by newspaper advertisements or from collaborating neurological clinics. The study was approved by the local ethical committee and all subjects gave written informed consent prior to inclusion.

2.2. Polysomnography

All subjects underwent overnight video-PSG with a SOMNOscreen™ plus Tele+Video (Somnomedics, Randersacker, Germany) device. Filter settings as well as scoring of sleep stages, arousals, respiratory events, and periodic leg movements followed standard recommendations and criteria (Berry et al. , 2017). Additionally, artifacts on the submental EMG were marked in the PSG software, i.e. transmission errors of the wireless PSG system or snoring artifacts with a typical waxing and waning pattern synchronous with airflow. RBD was diagnosed according to the criteria from the International Classification of Sleep Disorders version 3 (ICSD-3): A. Repeated episodes of sleep related vocalization and/or complex motor behaviors; B. These behaviors are documented by polysomnography to occur during REM sleep or, based on clinical history of dream enactment, are presumed to occur during REM sleep; C. Polysomnographic recording demonstrates REM sleep without atonia; D. The disturbance is not explained more clearly by another sleep disorder, mental disorder, medication, or substance use. All criteria had to be fulfilled. If

subjects had increased apnea-hypopnea indices (AHI) >5/h, dream enacting behavior had to be present in the absence of any flow events or corresponding arousals.

2.3. Visual scoring of REM sleep without atonia

Using DOMINO software (SOMNOmedics, Randersacker, Germany), two neurologists (MS, MO), board-certified in sleep medicine, visually scored RSWA in the first consecutive 10 PD^{RBD+} and 10 PD^{RBD-} subjects according to previously proposed criteria by McCarter and colleagues (McCarter et al. , 2014). Raters were blinded to each other and to the results of the computer algorithm. RSWA scoring was performed on the submental muscle alone. Background EMG activity was set as a period of 30 - 60s of REM sleep without any visible muscle activity, and the root mean square (RMS) of the EMG signal during this time was computed; baseline values ranged from 1.01 - 3.46 µV (mean 1.45 ± 0.52). Details of scoring as described by McCarter and colleagues are given in the appendix.

2.4. Algorithm for automatic computer detection of EMG activity

EMG recordings (in European data format, EDF) and visual user markings of REM sleep epochs and artifacts (flow events, arousals, and artifacts) were extracted from the DOMINO software. They were further processed and analyzed in a custom-made computer program using LabVIEW 2016 (National Instruments). Scoring guidelines for RSWA were implemented as follows: To prepare computation of EMG samples of 5ms precision, the EMG signal, recorded with a sampling rate of 256Hz, was resampled to 200Hz using spline interpolation. First, to detect 'tonic' activity, all 5ms samples with activity >2x baseline were marked as positive samples (= containing muscle activity) by a threshold detection. Second, if any artifact occurred during the 30s epoch, the whole epoch was excluded. Third, the time-array was divided into intervals of 0.2s (40 samples) and a search algorithm marked each 0.2s period as a positive period, if any sample within the period was a positive sample. Finally, if a 30s REM epoch contained any continuous period >15s with positive samples (>75 x 0.2s positive samples), the REM epoch was marked as positive for 'tonic' activity.

To detect 'phasic' activity, first the threshold for detection of the algorithm was set at >4x baseline activity, except during 'tonic' activity where the detection threshold was set to 2x RMS of the same 'tonic' activity period. This way, the phasic detection threshold changed for every registered 'tonic' activity, but only during the exact time period of the tonic activity. Second, all 5ms samples above the 'phasic' threshold were marked as positive samples and all marked artifacts were excluded in the same way as in the 'tonic'

activity detection. Third, a search algorithm was applied to find gaps of 0.1 - 0.2s (20 - 40 samples) between positive markings, and all samples were marked as positive samples if the gap was between 0.1 - 0.2s. The positive marking was eliminated in gaps > 0.2s or < 0.1s. In this way, all bursts less than 0.1s were discarded, and whenever there were more than 0.2s between two positive markings, the 'phasic' burst was considered to be terminated. Each 30s REM period was split into ten 3s mini-epochs, and if any artifact occurred during a mini-epoch, the entire mini-epoch was excluded from the analysis. A mini-epoch was regarded positive for 'phasic' activity if one or more 'phasic' bursts were registered at any time within the mini-epoch. If 'phasic' activity surpassed the limit between two mini-epochs, all mini-epochs containing activity were scored as epochs with 'phasic' activity. 'Any' activity was registered when either 'phasic' or 'tonic' activity was present within a mini-epoch. Percentage of 'phasic' and 'any' activity was calculated by dividing the number of mini-epochs with activity by the total number of REM sleep mini-epochs and percentage of 'tonic' activity was calculated accordingly on 30s REM sleep epochs.

Additionally, the duration as well as maximum amplitude of each 'phasic' burst were computed. Durations of all 'phasic' bursts were averaged for each subject for further analysis. Maximum amplitudes were normalized to individual's background EMG and were averaged for each subject for further analysis.

2.5. Statistical analysis

We analyzed the data with the Statistical Package for the Social Sciences (SPSS) version 24. Group data are presented as mean ± standard deviation or as relative frequencies if not otherwise stated. Groups were compared using Student's t-test, Mann-Whitney test, Kruskal-Wallis test and chi-square test as appropriate; non-parametric correlations were interrogated with Spearman's rho. Normal distribution of data was assessed with the Shapiro-Wilk test, Q-Q plots, and box plots. Coefficient of determination (R^2) and the Pitman-Morgan test were used to compare inter-rater variability on a subject level. Cohens Kappa was used to compare inter-rater variability on a single epoch level. Receiver operating curves (ROC) were calculated for 'tonic', 'phasic' and 'any' activity as well as for duration and amplitude of the 'phasic' bouts. Areas under the curve (AUC) were calculated for each analysis and cut-off thresholds are given for highest combined sensitivity and specificity, i.e. diagnostic accuracy.

3. Results

3.1. Comparison of computer algorithm to visual scorings

The computer algorithm was compared to observations of two human raters using the first consecutively recruited 10 PD^{RBD-} and 10 PD^{RBD+} subjects. 3'175 30s epochs for 'tonic', and 33'576 3s epochs for 'phasic' and 'any' activity were considered after elimination of flow events, arousals, and artifacts (17% of 30s and 12% of 3s epochs were discarded). Clinical characteristics did not differ between groups (**Table 1**). At a group level, elevated muscle activity in the PD^{RBD+} group was captured by both raters and the algorithm at high significance levels (all $p<0.001$, except $p=0.005$ for 'tonic' activity for rater 2), and there was good agreement between the two visual scorings and the computer algorithm; only 'tonic' activity in the PD^{RBD+} group was scored higher by rater 2 (**Table 2**). At an individual subject level, ratings for muscle activity showed a high correlation between both human raters ('tonic' activity, $R^2=0.936$, 'phasic' activity, $R^2=0.877$, 'any' activity, $R^2=0.982$), and the algorithm exhibited high congruency to these ratings ('tonic' activity, $R^2=0.976$ & $R^2=0.972$; 'phasic' activity, $R^2=.979$ & $R^2=0.846$; 'any' activity, $R^2=0.989$ & $R^2=0.982$), see **Figure 1**. The Pitman-Morgan test revealed that the residuals of the regressions of each human rater with the computer algorithm were not greater than the residuals of the regressions of the human ratings for all three activities. Agreements of human and computer scorings on single epoch level were high for all activities ('tonic' activity both raters 97%, 'phasic' activity 95% & 93%, 'any' activity both raters 95%), comparable to the agreement between the human raters (94%, 87%, and 98%, respectively). Frequency-corrected Cohen's κ showed high agreement between human raters and the algorithm ('tonic' activity $\kappa=0.74$ & $\kappa=0.78$, 'phasic' activity $\kappa=0.80$ & $\kappa=0.76$, 'any' activity $\kappa=0.83$ & $\kappa=0.82$), which was similar to the κ values between the two human raters ('tonic' activity $\kappa=0.64$, 'phasic' activity $\kappa=0.79$, 'any' activity $\kappa=0.82$).

3.2. Discrimination between RBD and non-RBD subjects by the computer algorithm

Next, we interrogated all included subjects with the computer algorithm to test for group differences in amount of RSWA. Overall, 6'345 30s epochs, and 69'181 3s epochs were analyzed after elimination of flow events, arousals, and artifacts (19% of 30s and 12% of 3s epochs were discarded). Groups were comparable with regard to demographic and clinical characteristics (**Table 3)**. RBD+ subjects exhibited highly significant increased 'tonic', 'phasic', and 'any' muscle activity compared to RBD- subjects (all p-values <0.001) (**Table 3**). At an individual patient level, cut-off values with full differentiation between RBD+ and RBD- subjects were obtained for 'phasic' and 'any' activity (cut off values at >10% of REM sleep; range of RBD- subjects,

1.3 - 8.2% for 'phasic' and 'any'; range of RBD+ subjects, 13.2 - 64.5% for 'phasic' and 13.3 - 77.4% for 'any', respectively). 'Tonic' activity gave an AUC of 0.976 at a cut-off value of 1.2% 'tonic' activity (range of RBD- subjects, 0 - 1.18%; range of RBD+ subjects, 0 - 66.3%, but only one iRBD subject had less than 1.2%) (**Figure 2**). Of note, when using all epochs without removal of epochs containing flow events, arousals, and artifacts, precision of group separation was lower and cut-off values differed: AUC of 'tonic' activity = 0.940 (cut-off = 2.2%), 'phasic' activity = 0.975 (cut-off = 15%), and 'any' activity = 0.978 (cut-off = 15%).

3.3. Duration and amplitude of phasic EMG activity and RSWA related to reported symptom severity

RBD+ subjects had longer duration of 'phasic' EMG bouts than RBD- subjects ($p<0.001$), and similarly, had higher amplitude of 'phasic' EMG bouts ($p=0.008$). ROC analysis resulted in an AUC of 0.909 for bout duration, and 0.768 for bout amplitude. A cut-off value of 405ms had a sensitivity of 91.3% and specificity of 82.8% to discriminate between RBD-negative and RBD-positive subjects. The amplitude value of 10.4 times of baseline had a sensitivity of 78.3% and specificity of 69%, respectively.

Subjects with higher scores on the RBDSQ exhibited a trend towards a higher amount of RSWA as shown in **Figure 3**; however, correlations were only significant when groups were pooled in the analysis ('any', rho=0.522, $p<0.001$; 'phasic', rho=0.515, $p<0.001$; 'tonic', rho=0.568, $p<0.001$).

4. Discussion

We have presented a semi-automated computer algorithm for scoring and quantification of REM sleep without atonia, separated into 'tonic', 'phasic', and 'any' activity, which showed good agreement with human ratings on a group, subject, and single epoch level. Our computer algorithm displayed high diagnostic accuracy when detecting and separating subjects with RBD from subjects without RBD. RBD not only manifested as an excess of submental muscle activity during REM sleep, but also as prolonged muscle activity with increased amplitude. Self-awareness of RBD symptoms correlated with the amount of altered muscle activity.

4.1. Considerations on the computer algorithm

Our results are in line with growing positive experience on computerization of scoring of muscle activity during REM sleep. The ''REM atonia index'' (RAI) computer algorithm developed by Ferri and colleagues has found the most widespread application and showed good agreement with visual scoring of RSWA on individual subject level (Ferri et al. , 2008, Ferri et al. , 2014). However, the RAI algorithm uses different rules for scoring muscle activity than proposed visual rules and does not adhere to the general 30s epoch-, and 3s mini-epoch-schemes, so a comparison with human ratings on an epoch level is difficult (Ferri et al. , 2008, Kempfner et al. , 2010, Frauscher et al. , 2013). Other computer algorithms also experienced good agreement with visual scorings on the subject level, but their validation mostly lacked comparison on a single epoch level or different measures of quantification of muscle activity were obtained with each method (Mayer et al. , 2008, Kempfner et al. , 2010, Frauscher et al. , 2014). In this study, we aimed to closely integrate visual scoring rules to a computer algorithm to obtain a high level of comparability between both methods.

4.2. Classification of RBD, artifact management

Our algorithm combined with visual artifact management correctly classified RBD subjects using a threshold of >1.2% 'tonic' activity and >10% 'phasic' or 'any' activity, irrespective of co-occurrence of PD or not. Its strong discriminating power with relatively low thresholds for altered muscle activity was a surprising finding, since accuracy is better than reported in previous studies using submental EMG alone (Lapierre et al. , 1992, Ferri et al. , 2008, Frauscher et al. , 2008, Ferri et al. , 2010, Kempfner et al. , 2010, Montplaisir et al. , 2010, Frauscher et al. , 2012, Frauscher et al. , 2013, McCarter et al. , 2014, Figorilli et

al. , 2017, McCarter et al. , 2017). We excluded REM sleep epochs that had a high likelihood of containing non-specific muscle activity, namely epochs containing apnea or hypopnea events and corresponding arousals (Iranzo et al. , 2005). We also visually assessed the submental EMG for artifacts during REM sleep; these were typically technical artifacts occurring because of brief transmission interruptions of the wireless PSG system as well as snoring artifacts with a typical waxing and waning pattern in the submental EMG synchronous to breathing or microphone measurements. Overall, 19% of 30s epochs and 12% of 3s epochs were excluded in our analysis; however, time consumption was low as assessment of breathing events and artifacts are part of a routine evaluation of a PSG. These numbers were higher than the ones reported by Figorilli et al. (Figorilli et al. , 2017), but could be one explanation for the high discriminative performance in our study. Without any artifact management, the discrimination power of our algorithm became slightly lower, but had similar AUC and cut-off values to those reported in its first description (McCarter et al. , 2014).

4.3. Duration and amplitude of phasic muscle activity

Our EMG analysis not only confirmed quantitatively increased muscle activity in RBD subjects, but also revealed increased duration and amplitude of 'phasic' EMG bouts. This is in line with the findings from McCarter and colleagues (McCarter et al. , 2014, McCarter et al. , 2017) as well as a previous report from Mayer et al (Mayer et al. , 2008). The etiology of RSWA genesis in humans is incompletely understood; based on animal models and scarce human studies, centers in the dorsal brainstem are suggested to play a key role in REM sleep initiation and maintenance (Lu et al. , 2006, Boeve et al. , 2007, Mayer et al. , 2015), and specifically the sublaterodorsal tegmental nucleus is supposed to be the key regulator of muscle atonia during REM sleep (Luppi et al. , 2011, Mayer et al. , 2015). Hence, malfunction of this complex in RBD might not only lead to a quantitative change of muscle atonia but also to a qualitative change of "leaked" muscle activity during REM sleep. Measures of duration and amplitude of 'phasic' activity might be useful for classification of borderline or unclear cases in a clinical setting.

4.4. Amount of RSWA as a severity index

We saw a tendency towards a higher degree of RSWA being linked to self-awareness of RBD symptoms as assessed with the RBDSQ. The RBDSQ was introduced as a screening questionnaire to identify subjects with RBD by asking for typical hallmarks of RBD expression (Stiasny-Kolster et al. , 2007). Even though it was not primarily designed to quantify the severity of RBD, one might argue that higher propensity to

dream enactment and classical RBD-like dream content might contribute to higher self-awareness and ultimately, represent a higher RBD severity. Additionally, a high level of RSWA was shown to predict the conversion of iRBD to PD in a longitudinal study (Postuma et al. , 2010), and was linked to more pronounced loss of dopaminergic terminals in a cross-sectional study (Eisensehr et al. , 2003). Additionally, a recent study reported a correlation between the amount of REM sleep atonia and the neuromelanin signal of the subcoeruleus/coeruleus complex in iRBD subjects and PD patients (Garcia-Lorenzo et al. , 2013, Ehrminger et al. , 2016) .Hence, the amount of RSWA may help not only to identify RBD subjects, but might also represent a marker of RBD severity. Wider accessibility of fast and easy-to-use computer solutions for assessment of RSWA amount might facilitate future research in this field (Boeve et al. , 2016).

4.5. Limitations of the study

Several limitations of the study have to be noted. First, we only analyzed a limited number of 52 subjects; specifically, our iRBD sample was restricted. Secondly, subjects were all recorded with a single PSG monitor at a single center, potentially introducing a selection and analysis bias. However, overall number of analyzed subjects were in a similar range to previous reports (Frauscher et al. , 2008, Frauscher et al. , 2012, McCarter et al. , 2014, Figorilli et al. , 2017, McCarter et al. , 2017), and we obtained similar values for 'tonic', 'phasic', and 'any' activity as reported in the initial report from McCarter and colleagues who used a different PSG monitor (McCarter et al. , 2014). We only included submental EMG recordings in our analysis, and inclusion of flexor digitorum superficialis, and extensor digitorum brevis muscles were reported to increase diagnostic accuracy for detection of RBD (Frauscher et al. , 2008). Even though, inclusion of these channels are preferable, they are not supported by all PSG monitors. On the other hand, assessment of qualitative changes of submental EMG activity, i.e. duration and amplitude of muscle bouts, and careful artifact management might overcome shortcomings from missing inclusion of additional muscles.

4.6. Conclusions

Our proposed semi-automated algorithm allows a fast quantitation of 'tonic', 'phasic', and 'any' activity with high agreement to visual ratings. Inclusion of a comprehensive visual artifact management may increase diagnostic accuracy by quantitative measures of altered muscle activity without significant time-consumption. Prolonged muscle activity with increased amplitude is part of the altered atonia during REM

sleep in RBD. Finally, our results suggest that quantitative measures of RSWA reflect symptom severity, which could be advantageous for clinical trials.

Role of funding source

The funding sources had no influence on the study design, collection and analysis of data, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Acknowledgments

This project received financial support from the Lundbeck Foundation, the Danish Association of Parkinson's disease (Parkinson-foreningen), the Swiss National Science Foundation (grant number, P2SKP3_161812), and the Hildegard Henssler-Stiftung.

Conflict of interest: The authors declare no conflicts of interests with regards to this manuscript.

A.1. Appendix

Elevated muscle activity was scored as ‘tonic’, ‘phasic’, and ‘any’ activity as follows: ‘tonic’ activity was scored on 30s epochs and a 30s epoch was defined positive for ‘tonic’ activity if >50% of the epoch had continuous EMG activity greater than double the background activity. The return of muscle activity to baseline levels for at least 0.2s was considered as the end of the ‘tonic’ activity. Epochs with flow events (apnea or hypopnea), arousals, or artifacts were discarded from analysis. Percentage of ‘tonic’ activity was calculated by dividing the number of 30s epochs with ‘tonic’ activity by the total number of REM sleep epochs. To score ‘phasic’, and ‘any’ activity, each 30s epoch was subdivided into 10 3s mini-epochs for analysis. ‘Phasic’ activity was defined as EMG activity with amplitude >4x of the background activity and duration from 0.1 - 14.9s. The return of muscle activity to baseline levels for at least 0.2s was considered as end of a phasic burst. If ‘phasic’ activity surpassed the limit between two mini-epochs within a 30s epoch, all mini-epochs containing activity were scored as epochs with ‘phasic’ activity. During an epoch with ‘tonic’ activity, ‘phasic’ activity was scored, if EMG activity was >2x of the new baseline during the ‘tonic’ activity (these epochs were counted as positive for ‘tonic’ and ‘phasic’) according to the proposed rules by McCarter and colleagues (McCarter et al. , 2014). Mini-epochs containing flow events, arousals, or artifacts were discarded from analysis. Percentage of ‘phasic’ activity was calculated by dividing the number of 3s mini-epochs with ‘phasic’ activity by the total number of 3s REM sleep mini-epochs. ‘Any’ activity was defined as 3s mini-epochs containing either ‘tonic’ or ‘phasic’ or both activities.

References

Berry RB, Brooks R, Gamaldo C, Harding SM, Lloyd RM, Quan SF, et al. AASM Scoring Manual Updates for 2017 (Version 2.4). J Clin Sleep Med. 2017;13:665-6.

Boeve BF, Silber MH, Saper CB, Ferman TJ, Dickson DW, Parisi JE, et al. Pathophysiology of REM sleep behaviour disorder and relevance to neurodegenerative disease. Brain. 2007;130:2770-88.

Boeve BF, St Louis EK, Kantarci K. Neuromelanin-sensitive imaging in patients with idiopathic rapid eye movement sleep behaviour disorder. Brain. 2016;139:1005-7.

Chahine LM, Xie SX, Simuni T, Tran B, Postuma R, Amara A, et al. Longitudinal changes in cognition in early Parkinson's disease patients with REM sleep behavior disorder. Parkinsonism Relat Disord. 2016;27:102-6.

Ehrminger M, Latimier A, Pyatigorskaya N, Garcia-Lorenzo D, Leu-Semenescu S, Vidailhet M, et al. The coeruleus/subcoeruleus complex in idiopathic rapid eye movement sleep behaviour disorder. Brain. 2016;139:1180-8.

Eisensehr I, Linke R, Tatsch K, Kharraz B, Gildehaus JF, Wetter CT, et al. Increased muscle activity during rapid eye movement sleep correlates with decrease of striatal presynaptic dopamine transporters. IPT and IBZM SPECT imaging in subclinical and clinically manifest idiopathic REM sleep behavior disorder, Parkinson's disease, and controls. Sleep. 2003;26:507-12.

Ferri R, Gagnon JF, Postuma RB, Rundo F, Montplaisir JY. Comparison between an automatic and a visual scoring method of the chin muscle tone during rapid eye movement sleep. Sleep Med. 2014;15:661-5.

Ferri R, Manconi M, Plazzi G, Bruni O, Vandi S, Montagna P, et al. A quantitative statistical analysis of the submentalis muscle EMG amplitude during sleep in normal controls and patients with REM sleep behavior disorder. J Sleep Res. 2008;17:89-100.

Ferri R, Rundo F, Manconi M, Plazzi G, Bruni O, Oldani A, et al. Improved computation of the atonia index in normal controls and patients with REM sleep behavior disorder. Sleep Med. 2010;11:947-9.

Figorilli M, Ferri R, Zibetti M, Beudin P, Puligheddu M, Lopiano L, et al. Comparison Between Automatic and Visual Scorings of REM Sleep Without Atonia for the Diagnosis of REM Sleep Behavior Disorder in Parkinson Disease. Sleep. 2017;40.

Frauscher B, Ehrmann L, Hogl B. Defining muscle activities for assessment of rapid eye movement sleep behavior disorder: from a qualitative to a quantitative diagnostic level. Sleep Med. 2013;14:729-33.

Frauscher B, Gabelia D, Biermayr M, Stefani A, Hackner H, Mitterling T, et al. Validation of an integrated software for the detection of rapid eye movement sleep behavior disorder. Sleep. 2014;37:1663-71.

Frauscher B, Iranzo A, Gaig C, Gschliesser V, Guaita M, Raffelseder V, et al. Normative EMG values during REM sleep for the diagnosis of REM sleep behavior disorder. Sleep. 2012;35:835-47.

Frauscher B, Iranzo A, Hogl B, Casanova-Molla J, Salamero M, Gschliesser V, et al. Quantification of electromyographic activity during REM sleep in multiple muscles in REM sleep behavior disorder. Sleep. 2008;31:724-31.

Garcia-Lorenzo D, Longo-Dos Santos C, Ewenczyk C, Leu-Semenescu S, Gallea C, Quattrocchi G, et al. The coeruleus/subcoeruleus complex in rapid eye movement sleep behaviour disorders in Parkinson's disease. Brain. 2013;136:2120-9.

Iranzo A, Santamaria J. Severe obstructive sleep apnea/hypopnea mimicking REM sleep behavior disorder. Sleep. 2005;28:203-6.

Kempfner J, Sorensen G, Zoetmulder M, Jennum P, Sorensen HB. REM behaviour disorder detection associated with neurodegenerative diseases. Conf Proc IEEE Eng Med Biol Soc. 2010;2010:5093-6.

Lapierre O, Montplaisir J. Polysomnographic features of REM sleep behavior disorder: development of a scoring method. Neurology. 1992;42:1371-4.

Lu J, Sherman D, Devor M, Saper CB. A putative flip-flop switch for control of REM sleep. Nature. 2006;441:589-94.

Luppi PH, Clement O, Sapin E, Gervasoni D, Peyron C, Leger L, et al. The neuronal network responsible for paradoxical sleep and its dysfunctions causing narcolepsy and rapid eye movement (REM) behavior disorder. Sleep Med Rev. 2011;15:153-63.

Mayer G, Bitterlich M, Kuwert T, Ritt P, Stefan H. Ictal SPECT in patients with rapid eye movement sleep behaviour disorder. Brain. 2015;138:1263-70.

Mayer G, Kesper K, Ploch T, Canisius S, Penzel T, Oertel W, et al. Quantification of tonic and phasic muscle activity in REM sleep behavior disorder. J Clin Neurophysiol. 2008;25:48-55.

McCarter SJ, St Louis EK, Duwell EJ, Timm PC, Sandness DJ, Boeve BF, et al. Diagnostic thresholds for quantitative REM sleep phasic burst duration, phasic and tonic muscle activity, and REM atonia index in REM sleep behavior disorder with and without comorbid obstructive sleep apnea. Sleep. 2014;37:1649-62.

McCarter SJ, St Louis EK, Sandness DJ, Duwell EJ, Timm PC, Boeve BF, et al. Diagnostic REM sleep muscle activity thresholds in patients with idiopathic REM sleep behavior disorder with and without obstructive sleep apnea. Sleep Med. 2017;33:23-9.

Montplaisir J, Gagnon JF, Fantini ML, Postuma RB, Dauvilliers Y, Desautels A, et al. Polysomnographic diagnosis of idiopathic REM sleep behavior disorder. Mov Disord. 2010;25:2044-51.

Neikrug AB, Avanzino JA, Liu L, Maglione JE, Natarajan L, Corey-Bloom J, et al. Parkinson's disease and REM sleep behavior disorder result in increased non-motor symptoms. Sleep Med. 2014;15:959-66.

Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2015a;30:1591-601.

Postuma RB, Gagnon JF, Bertrand JA, Genier Marchand D, Montplaisir JY. Parkinson risk in idiopathic REM sleep behavior disorder: preparing for neuroprotective trials. Neurology. 2015b;84:1104-13.

Postuma RB, Gagnon JF, Rompre S, Montplaisir JY. Severity of REM atonia loss in idiopathic REM sleep behavior disorder predicts Parkinson disease. Neurology. 2010;74:239-44.

St Louis EK, Boeve AR, Boeve BF. REM Sleep Behavior Disorder in Parkinson's Disease and Other Synucleinopathies. Mov Disord. 2017;32:645-58.

Stiasny-Kolster K, Mayer G, Schafer S, Moller JC, Heinzel-Gutenbrunner M, Oertel WH. The REM sleep behavior disorder screening questionnaire--a new diagnostic instrument. Mov Disord. 2007;22:2386-93.

Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord. 2010;25:2649-53.

Tables and Figures

Table 1: Demographic and clinical characteristics of validation cohort

	PD^{RBD-}	PD^{RBD+}	P-value
Age	63.5 ± 9.2	66.4 ± 10.2	ns‡
Sex (male / female)	6 / 4	8 / 2	ns§
Disease duration [y]	5.0 ± 3.9	7.5 ± 4.3	ns¤
Hoehn & Yahr	2.0 ± 0.6	2.4 ± 0.6	ns¤
MDS-UPDRS III, total	30.9 ± 11.7	37.8 ± 12.2	ns‡
LED [mg]	576.3 ± 523.3	817.8 ± 307.0	ns‡
Sleep characteristics			
RBDSQ	3.6 ± 2.3	6.7 ± 2.4	0.019¤
Sleep efficiency [%]	86.6 ± 6.1	78.0 ± 12.8	ns¤
Sleep latency [min]	16.8 ± 20.6	16.0 ± 11.4	ns¤
Wake [%]	13.4 ± 6.0	22.0 ± 12.8	ns¤
N1 [%]	12.4 ± 6.6	11.6 ± 6.1	ns‡
N2 [%]	36.1 ± 8.6	38.9 ± 7.7	ns‡
SWS [%]	15.8 ± 8.2	9.4 ± 3.1	ns‡
REM [%]	22.2 ± 9.1	18.0 ± 10.6	ns¤
AHI [h^{-1}]	6.7 ± 6.5	8.2 ± 7.0	ns¤

‡ = parametric test (Student's t-test), ¤ = non-parametric test (Mann-Whitney test), § = χ^2 test.

Abbreviations: AHI, apnea hypopnea index, LED, levodopa equivalent dose, N1, non-REM sleep 1, N2, non-REM sleep 2, ns, not significant, RBD, rapid eye movement sleep behavior disorder, RBDSQ, RBD screening questionnaire, REM, rapid eye movement sleep, PD, Parkinson's disease

Table 2: REM sleep atonia measures between raters and algorithm

	PD^{RBD-}			P-Value	PD^{RBD+}			P-value
RSWA indices	*algorithm*	*rater1*	*rater2*		*algorithm*	*rater1*	*rater2*	
Tonic [%]	0.1 ± 0.3	0.1 ± 0.3	1.3 ± 2.3	ns¤	15.2 ± 18.4	13.2 ± 16.3	17.7 ± 20.9	0.030¤ *
Phasic [%]	3.9 ± 1.7	3.8 ± 1.8	4.2 ± 2.2	ns¤	31.0 ± 13.8	30.9 ± 15.9	32.5 ± 16.7	ns¤
Any [%]	3.9 ± 1.7	3.8 ± 1.8	3.2 ± 2.8	ns¤	34.7 ± 17.9	34.2 ± 18.4	39.5 ± 20.4	ns¤

¤ = non-parametric test (Kruskal-Wallis test for all three groups, and Mann-Whitney test for two groups, respectively), *Pair-wise comparison: rater2 against rater1, p=0.011, and rater2 against algorithm, p=0.043; rater1 against algorithm not significant.

Abbreviations: ns, not significant, RBD, rapid eye movement sleep behavior disorder PD, Parkinson's disease

Table 3: Demographic and clinical characteristics of groups analyzed with the algorithm

	RBD- subjects			**RBD+ subjects**			**P-value**
	all	HC	PD^{RBD-}	**all**	PD^{RBD+}	iRBD	
Age	**66.2 ± 7.4**	66.9 ± 5.8	65.4 ± 9.0	**66.4 ± 8.2**	66.7 ± 9.7	65.7 ± 3.9	ns‡
Sex (male / female)	**20 / 9**	10 / 5	10 / 4	**18 / 5**	12 / 4	6 / 1	ns§
Disease duration [y]			5.0 ± 3.6		8.0 ± 4.4		0.045¤
Hoehn & Yahr			2.1 ± 0.6		2.3 ± 0.5		ns¤
MDS-UPDRS III, total			32.6 ± 11.9		38.9 ± 10.5		ns‡
LED [mg]			547.0 ± 467.6		747.2 ± 329.6		ns‡
Sleep characteristics							
RBDSQ	**3.0 ± 1.9**	2.8 ± 1.7	3.3 ± 2.2	**8.0 ± 3.2**	7.3 ± 3.4	10.0 ± 1.7	<0.001¤
Sleep efficiency [%]	**84.7 ± 6.7**	83.5 ± 5.9	86.1 ± 5.9	**84.3 ± 11.4**	80.9 ± 12.0[1]	92.1 ± 3.9	ns¤
Sleep latency [min]	**14.8 ± 13.9**	13.3 ± 7.9	16.4 ± 18.6	**16.4 ± 11.3**	15.0 ± 11.4	19.6 ± 11.2	ns¤
Wake [%]	**15.2 ± 6.7**	16.5 ± 7.4	13.9 ± 5.8	**15.7 ± 11.4**	19.1 ± 12.0[1]	7.9 ± 3.9	ns¤
N1 [%]	**15.3 ± 6.4**	17.5 ± 5.4	13.0 ± 6.8	**13.7 ± 6.6**	12.6 ± 6.1	16.3 ± 7.5	ns‡
N2 [%]	**38.3 ± 10.4**	38.0 ± 11.3	38.7 ± 9.7	**43.6 ± 9.5**	41.7 ± 8.7	48.0 ± 10.3	ns‡
SWS [%]	**14.1 ± 7.0**	13.3 ± 5.9	14.9 ± 8.1	**11.3 ± 4.2**	10.3 ± 3.2	13.8 ± 5.8	ns‡
REM [%]	**15.7 ± 8.2**	12.6 ± 4.6	19.1 ± 9.9	**15.5 ± 8.2**	16.2 ± 9.3	13.9 ± 5.1	ns¤
AHI [h^{-1}]	**10.6 ± 8.5**	12.7 ± 9.0	8.3 ± 7.5	**9.4 ± 10.8**	7.2 ± 6.2	15.3 ± 17.9	ns¤
RSWA indices							
Tonic [%]	**0.1 ± 0.3**	0.1 ± 0.3	0.1 ± 0.2	**17.3 ± 19.8**	19.4 ± 21.5	12.7 ± 15.8	<0.001¤
Phasic [%]	**3.7 ± 2.0**	3.2 ± 2.1	4.2 ± 1.8	**34.2 ± 13.5**	33.6 ± 12.6	35.6 ± 16.4	<0.001¤
Any [%]	**3.7 ± 2.0**	3.2 ± 2.1	4.2 ± 1.8	**38.5 ± 17.7**	38.5 ± 18.3	38.6 ± 17.8	<0.001¤
Duration [ms]	**338.0 ± 77.8**	328.8 ± 91.9	348.0 ± 49.1	**487.7 ± 77.3**	489.0 ± 82.3	484.7 ± 70.4	<0.001‡
Max amplitude	**10.2 ± 3.2**	9.7 ± 3.5	10.8 ± 2.9	**12.4 ± 2.4**	12.4 ± 2.5	12.4 ± 2.2	0.008‡

‡ = parametric test (Student's t-test), ¤ = non-parametric test (Mann-Whitney test), § = χ^2 test, [1] = PD^{RBD+} and iRBD differed significantly

Abbreviations: AHI, apnea hypopnea index, HC, healthy control subjects, iRBD, idiopathic RBD subjects, LED, levodopa equivalent dose, N1, non-REM sleep 1, N2, non-REM sleep 2, RBD, rapid eye movement sleep behavior disorder, RBDSQ, RBD screening questionnaire, REM, rapid eye movement sleep, PD, Parkinson's disease

Figure 1: Validation of computer algorithm against raters on single subject level

Figure A: Correlation plot of human raters assessing percentage of 'tonic', 'phasic', and 'any' muscle activity during REM sleep; line of unity in light grey. **Figure B:** Correlation plot of computer algorithm scoring of percentage of 'tonic', 'phasic', and 'any' muscle activity during REM sleep against the two human raters from A; line of unity in light grey.

Figure 2: Amount of increased muscle activity during REM sleep

Dot-plots of single subjects' percentage of increased 'tonic', 'phasic' and 'any' muscle activity during REM sleep. RBD+ subjects (PD^{RBD+} and iRBD) exhibited increased activity in all three qualities with cut-off values allowing best separation of groups of 1.2% for 'tonic' (line not shown), and 10% for 'phasic' and 'any' activity (light grey line), respectively. Note that only three RBD- subjects (one PD^{RBD-} and two HC) had tonic activity, and only one RBD-positive subject had no 'tonic' activity.

Abbreviations: HC, healthy control subjects, iRBD, idiopathic RBD subjects, RBD, rapid eye movement sleep behavior disorder, PD, Parkinson's disease

Figure 3: Correlation of RBDSQ and REM sleep without atonia

Correlation plot of score on the rapid eye movement (REM) sleep behavior disorder screening questionnaire (RBDSQ) and percentage of increased 'any' activity during REM sleep.

Abbreviations: HC, healthy control subjects, iRBD, idiopathic RBD subjects, RBD, rapid eye movement sleep behavior disorder, PD, Parkinson's disease